3.0 Validation Status of the rLLNA Test Method

The following is a synopsis of the information in the final ICCVAM BRD (**Appendix D**), which reviews the available data and information for the rLLNA test method. The ICCVAM BRD describes the current validation status of the rLLNA test method, including what is known about its reliability and accuracy, the scope of the substances tested, and standardized protocols used for the validation study.

3.1 Test Method Description

The purpose of the rLLNA test method is to identify potential skin sensitizers by quantifying lymphocyte proliferation. The mechanistic basis is identical to that of the traditional LLNA, which measures the magnitude of lymphocyte proliferation, which in turn correlates with the extent to which sensitization develops after a topical induction exposure to a skin-sensitizing substance.

With one exception, the technical aspects of the rLLNA are identical to those of the traditional LLNA (ICCVAM 1999). The traditional LLNA tests three dose levels of each test substance for skin-sensitizing activity. In the rLLNA, only one dose of the test substance is tested: the concentration that provides maximum solubility without causing overt systemic toxicity and/or excessive skin irritation (Kimber et al. 2006). Guidance for evaluating local irritation and systemic toxicity in the LLNA is provided in the updated ICCVAM-recommended LLNA protocol (**Appendix B**).

3.1.1 General Test Method Procedures

The rLLNA measures lymphocyte proliferation after topical exposure to a potential skin-sensitizing substance. The test substance is administered topically on three consecutive days to the ears of mice at a concentration that provides maximum solubility of the test substance without systemic toxicity and/or excessive local irritation. Two days after the final application of the test substance, ^{3}H-thymidine or ^{125}I-iododeoxyuridine (in phosphate-buffered saline; 250 µL/mouse) is administered via the tail vein. Five hours later the draining auricular lymph nodes are excised, and a single-cell suspension from the lymph nodes of each animal is prepared for quantifying the incorporation of radioactivity, which correlates with lymph node cell proliferation.

The incorporation of radioactive ^{3}H-thymidine or ^{125}I-iododeoxyuridine for each mouse is expressed in disintegrations per minute (dpm). The SI is calculated as the ratio of the mean dpm/mouse for each treatment group against the mean dpm/mouse for the vehicle control group. The threshold for a positive response is an SI ≥ 3.

3.1.2 Similarities and Differences between the Protocols for the Traditional LLNA and the rLLNA

As mentioned above, the only difference between the traditional LLNA (ICCVAM 1999) and the rLLNA is that only one test substance dose is included in rLLNA, while three doses are tested in the traditional LLNA. All other procedures are identical.

3.2 Validation Database

Data were obtained from 11 different sources, including published reports and unpublished data submitted to NICEATM in response to a May 17, 2007, *FR* notice (72 FR 27815[19]). The rLLNA database consisted of the results for the highest doses tested in these studies.

The resulting database consisted of 457 unique substances tested in a total of 471 traditional LLNA studies (**Table 3-1**), 211 of which were included in the original ICCVAM evaluation of the traditional LLNA (ICCVAM 1999). Fourteen of the 457 unique substances[20] were repeated from two to five times in different LLNA studies. Specifically, nine of the 14 substances were evaluated two to five times in different vehicles, and five of the 14 substances were evaluated two to five times in the same vehicle. Two of the five substances evaluated in the same vehicle (hexyl cinnamic aldehyde [HCA] and potassium dichromate) were also tested using different vehicles (one study for HCA and two studies for potassium dichromate). Due to the small number of repeated studies (5% of total studies), all studies were treated independently for the purpose of this accuracy evaluation. When the studies for the substances repeated in the same vehicle were considered together to yield an overall skin sensitization classification, there were 465 studies with unique substance and vehicle combinations.

Table 3-1 provides the chemical class information for these test substances. The table distinguishes the chemical classifications of the 211 substances included in the original evaluation of the rLLNA (Kimber et al. 2006; ESAC 2007) and the chemical classifications of the additional substances received in response to the *FR* notice. Of the 211 substances initially evaluated by Kimber et al. (2006), the chemical classes with the greatest number of substances were carboxylic acids (29) and halogenated hydrocarbons (27). Of the additional 246 substances included in this evaluation, the chemical classes with the greatest number of substances tested were pharmaceutical chemicals (125), carboxylic acids (15), and lipids (14). Of the substances included in this evaluation, 10 were formulations. Seventy substances could not be assigned to a specific chemical class due to incomplete available information (e.g., the lack of a Chemical Abstracts Service Registry Number or structure).

[19] Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf

[20] Some substances were tested in more than one vehicle. In such instances, each substance–vehicle combination was considered separately, thus a total of 465 unique substance–vehicle combinations were used in the performance evaluation.

Table 3-1 Chemical Classes[1] Represented in the Current Traditional LLNA Database

Chemical Class	Number of Substances - Original[2]	Number of Substances - Additional[2]	Chemical Class	Number of Substances - Original	Number of Substances - Additional
Alcohols	9	4	Inorganic Chemicals	0	2
Aldehydes	21	4	Isocyanates	1	0
Amides	4	0	Ketones	5	0
Amidines	1	0	Lactones	2	2
Amines	14	7	Lipids	7	14
Anhydrides	1	0	Macromolecular substances[3]	0	5
Carbohydrates	3	2	Nitriles	1	1
Carboxylic acids	29	15	Nitro compounds	2	0
Esters	3	0	Nitroso compounds	3	0
Ethers	14	2	Onium compounds	1	0
Formulations[3]	0	10	Pharmaceutical chemicals[4]	0	125
Heterocyclic compounds	18	4	Phenols	18	2
Hydrocarbons, Acyclic	2	1	Polycyclic compounds	5	3
Hydrocarbons, Cyclic	14	7	Quinones	1	1
Hydrocarbons, halogenated	27	1	Sulfur compounds	20	2
Hydrocarbons, other	7	8	Urea	3	0
Imines	0	1	Unknown	28	42

1 Total number of substances assigned to chemical classes does not equal the total number of substances evaluated because some substances were assigned to more than one class and some substances were not assigned to a specific chemical class.

2 Number of substances - Original represents the substances evaluated in Kimber et al. (2006).
Number of substances - Additional represents the substances received in response to the released *Federal Register* notice (72 FR 27815, May 17, 2007) available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf.

3 No chemical class could be assigned. The terms "formulation" and "macromolecular substances" were used to classify these substances.

4 The chemical classification of "pharmaceutical chemicals" for the GlaxoSmithKline (GSK) substances was suggested by Dr. Michael Olson of GSK to capture three types of pharmaceutical substances (actives, intermediates, and starting materials).

3.3 Reference Test Method Data

The traditional LLNA data used for evaluation of the rLLNA include the results for all tested doses of each substance. In addition to calculated SI values for each of the tested doses, the vehicles tested and EC3 values (estimated concentration needed to produce an SI value of 3) for substances classified as sensitizers were provided in Gerberick et al. (2005). The data received in response to the May 2007 *FR* notice included calculated SI values for the vehicle and each of the tested doses. If EC3 values were not included in the data source, they were calculated, where possible, using either interpolation or extrapolation (Dearman et al. 2007). This information and the complete database (by each source) are provided in **Annex III** of the BRD (**Appendix D**).

3.4 Test Method Accuracy

The ability of the rLLNA to correctly identify potential skin sensitizers was compared to that of the traditional LLNA. Of the 471 studies, 318 detected skin sensitizers, and 153 detected non-sensitizers. When studies of the substances tested more than once in the same vehicle were considered together to yield an overall skin sensitization classification, 465 unique substance–vehicle combinations resulted. Of these, 315 were identified as sensitizers and 150 as non-sensitizers.

Based on the available study data, the rLLNA has an accuracy of 98.7% (465/471), a sensitivity of 98.1% (312/318), a specificity of 100% (153/153), a false positive rate of 0% (0/153), and a false negative rate of 1.9% (6/318) when compared to the traditional LLNA (**Table 3-2**). When substances tested more than once in the same vehicle were considered together, the resulting 465 studies had an accuracy of 98.7% (459/465), a sensitivity of 98.1% (309/315), a specificity of 100% (150/150), a false positive rate of 0% (0/150), and a false negative rate of 1.9% (6/315).

This analysis of the rLLNA yielded six false negative results. A review of the data for these six substances indicates that the traditional LLNA classification of the substances as skin sensitizers was based not on the highest tested dose but on a low- or mid-dose level that produced an SI ≥ 3, while the highest dose tested produced an SI < 3. Because the rLLNA tests substances at only the highest dose level, all six substances would be incorrectly identified as non-sensitizers (i.e., false negatives). Four of the six substances that resulted in false negatives using the rLLNA compared to the traditional LLNA came from LLNA studies that used pooled data. There were no patterns of consistency for these substances with regard to physicochemical properties.

Table 3-2 Evaluation of the Performance of the rLLNA in Predicting Skin Sensitizers Compared to the Traditional LLNA

Data	N	Accuracy	Sensitivity	Specificity	False Positive	False Negative
Kimber et al. (2006)	211	98.6% (208/211)	98.2% (166/169)	100% (42/42)	0% (0/42)	1.8% (3/169)
rLLNA	471	98.7% (465/471)	98.1% (312/318)	100% (153/153)	0% (0/153)	1.9% (6/318)
rLLNA approach (substances repeated in the same vehicle considered together)	465	98.7% (459/465)	98.1% (309/315)	100% (150/150)	0% (0/150)	1.9% (6/315)

Abbreviation: N = number of tests
Accuracy = the percentage of correct outcomes (positive and negative) of a test method
Sensitivity = the percentage of all positive substances that are classified as positive
Specificity = the percentage of all negative substances that are classified as negative
False positive rate = the percentage of all negative substances that are falsely identified as positive
False negative rate = the percentage of all positive substances that are falsely identified as negative

3.5 Test Method Reliability

The BRD assessed interlaboratory reproducibility of the rLLNA with traditional LLNA data for five substances that had been tested independently in the same vehicle at multiple laboratories. These five substances were dinitrochlorobenzene (DNCB), HCA, linalool alcohol, methyl salicylate, and potassium dichromate. **Table 3-3** summarizes the responses obtained by the rLLNA. All studies classified DNCB, methyl salicylate, and potassium dichromate (3/5 = 60%) as sensitizers or non-sensitizers (i.e., 100% concordance). HCA and linalool alcohol, which were tested independently in two laboratories, were each classifed as a sensitizer by one traditional LLNA study and as a non-sensitizer by the other traditional LLNA study. Review of the studies indicates that the discordant results were due to differences in the highest dose levels tested. However, because the rLLNA and traditional LLNA use identical protocols and the data sets used to evaluate their accuracy are similar, the intra- and interlaboratory reliability of the rLLNA is deemed to be similar to that of the traditional LLNA (see ICCVAM 1999 for these statistics).

Table 3-3 Interlaboratory Reproducibility of Skin sensitization Outcome for the rLLNA

Substance	Data Source	Vehicle	rLLNA Dose (%)/SI	rLLNA Classification[1]
1-Chloro-2-dinitrobenzene	Gerberick et al. (2005)	AOO	0.25/38.00	+
	Data submitted by D. Germolec		0.25/7.10	+
Hexyl cinnamic aldehyde	Gerberick et al. (2005)	AOO	50/17.00	+
	Data submitted by H.W. Vohr		10/2.84	–
Linalool alcohol	Gerberick et al. (2005)	AOO	100/8.30	+
	Data submitted by D. Basketter, I. Kimber, and F. Gerberick		30/1.30	–
Methyl salicylate	Gerberick et al. (2005)	AOO	20/0.90	–
	Data submitted by D. Germolec		20/1.72	–
Potassium dichromate	Gerberick et al. (2005)	DMSO	0.5/16.10	+
	Data submitted by D. Germolec		0.25/3.39	+
	Ryan et al. (2002)		0.5/10.10	+

Abbreviations: AOO = Acetone: olive oil; DMSO = Dimethyl sulfoxide; rLLNA = Reduced murine local lymph node assay; SI = stimulation index

[1] – = non-sensitizer, + = sensitizer

3.6 Animal Welfare Considerations: Reduction, Refinement, and Replacement

Compared to the traditional LLNA, the rLLNA will reduce the number of animals used to assess skin sensitization. Becuse the rLLNA tests only the highest dose level of the test substance in addition to the concurrent control groups, the number of animals tested would be decreased by at least 40% for each test. Ryan et al. (2008) described the impact of reducing the number of animals per group from five to two on the performance of the rLLNA and concluded that such a small number of animals per group was inadequate for hazard identification of skin sensitizers.